

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.

PAT-NO: JP407271706A
DOCUMENT-IDENTIFIER: JP 07271706 A
TITLE: FAULT RESISTANT COMMAND PROCESSING SYSTEM AND METHOD AND RELATING DEVICE
PUBN-DATE: October 20, 1995

INVENTOR-INFORMATION:

NAME	COUNTRY
CARLSON, WAYNE C	
MERRILL, BURNS GRECO	
MICHAEL, J MEADOWS	

ASSIGNEE-INFORMATION:

NAME	COUNTRY
INTERNATL BUSINESS MACH CORP	N/A

APPL-NO: JP07007879
APPL-DATE: January 23, 1995

INT-CL (IPC): G06F013/12 , G06F011/20 , G06F015/16

ABSTRACT:

PURPOSE: To provide a fault resistant redundant processing system provided with efficient inter-unit communication by specifying a first command processing unit as an active unit and specifying a second command processing unit as a standby unit.

CONSTITUTION: A redundant command processor validates a command not considered as valid even though it is received regardless of the fault of one processor, prevents the command queued for execution from being lost, completes the command of a process during execution and transmits a response not transmitted yet. The command is received and made to standby in a queue by both active unit 12 and standby unit 14 but the command is processed only by the active unit 12. In the case that the active unit 12 fails, the standby unit 14 is automatically turned to be on-line, processes the unprocessed command and transmits the response to a command source 16. When a fault unit is restored, the unit is turned to the standby unit.

COPYRIGHT: (C)1995,JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-271706

(43) 公開日 平成7年(1995)10月20日

(51) Int. Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 13/12	3 4 0 B	7368-5B		
11/20	3 1 0 A			
15/16	4 7 0 B			

審査請求 有 請求項の数24 O L (全 12 頁)

(21) 出願番号 特願平7-7879

(22) 出願日 平成7年(1995)1月23日

(31) 優先権主張番号 2 1 7 8 2 9

(32) 優先日 1994年3月25日

(33) 優先権主張国 米国 (US)

(71) 出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション

INTERNATIONAL BUSINESS MACHINES CORPORATION

アメリカ合衆国10504、ニューヨーク州アーモンク (番地なし)

(72) 発明者 ウェイン・チャールズ・カールソン

アメリカ合衆国85715、アリゾナ州ツーソン、イースト・フェーンヒル・ドライブ8480

(74) 代理人 弁理士 合田 潔 (外2名)

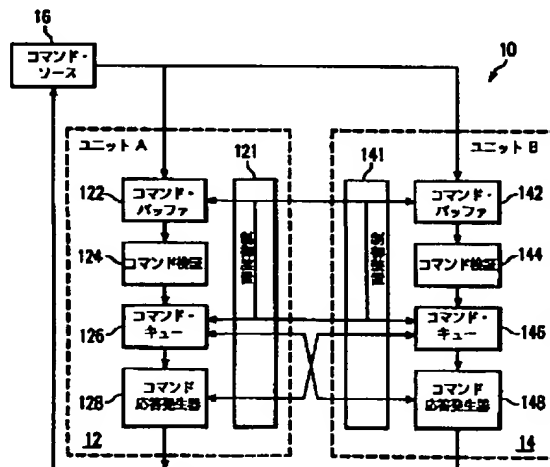
最終頁に続く

(54) 【発明の名称】 耐障害コマンド処理システム及び方法並びに関連する装置

(57) 【要約】

【目的】 冗長プロセッサにより、従来の耐障害処理に要求されるオーバヘッドを低減する方法及び装置を提供する。

【構成】 コマンドがアクティブ状態及び待機状態の両ユニットにより受信され、キューに待機されるが、アクティブ・ユニットによってのみ処理される。各々がコマンド識別子を含む短いメッセージが、アクティブ・ユニットから待機ユニットへ送信される。更に、周期的なハンドシェーキング信号が2つのユニット間で交換される。アクティブ・ユニットが故障すると、待機ユニットが自動的にオンラインとなり、未処理のコマンドを処理し、応答をコマンド・ソースに送信する。故障ユニットが復元すると、今度はそれが待機ユニットとなる。



1

【特許請求の範囲】

【請求項1】耐障害コマンド処理のためのシステムであって、

コマンド・ソースからコマンドを受信する第1の受信機と、

前記第1の受信機に結合され、前記コマンド・ソースに
 応答を送信する第1の送信機と、

前記第1の受信機及び前記第1の送信機に結合される第
 1のユニット間インタフェースと、

を含む、第1のコマンド処理ユニットと、

前記コマンド・ソースからコマンドを受信する第2の受
 信機と、

前記第2の受信機に結合され、前記コマンド・ソースに
 応答を送信する第2の送信機と、

前記第2の受信機、前記第2の送信機及び前記第1のユ
 ニット間インタフェースに結合される第2のユニット間
 インタフェースと、

を含む、第2のコマンド処理ユニットと、

前記第1のコマンド処理ユニットをアクティブ・ユニッ
 トとして指定し、前記第2のコマンド処理ユニットを待
 機ユニットとして指定する指定モジュールと、

を含み、前記第1の受信機及び前記第1の送信機が前記
 第2の受信機に動作的にリンクされる、システム。

【請求項2】コマンド・ソースからのコマンドが第2の
 受信機により受信されない場合に、該コマンドを前記第
 1の受信機から前記第2の受信機に送信するように指令
 する手段を含む、請求項1記載のシステム。

【請求項3】前記指定モジュールが、
 前記第1のコマンド処理ユニットが故障を発生したことを
 検出する手段と、

前記第2のコマンド処理ユニットをアクティブ・ユニッ
 トとして再指定する手段と、

を含む、請求項1記載のシステム。

【請求項4】前記指定モジュールが、
 前記第1のコマンド処理ユニットが復元されたことを検
 出する手段と、

前記第1のコマンド処理ユニットを待機ユニットとして
 再指定する手段と

を含み、前記第2の受信機及び前記第2の送信機が動作
 的に前記第1の受信機にリンクされる、請求項3記載の
 システム。

【請求項5】前記第1の受信機が、
 前記コマンド・ソースに結合される第1のバッファと、

前記第1のバッファからコマンドを受信するように結合
 される第1の検証モジュールと、

前記第1の検証モジュールから検証済みコマンドを受信
 するように結合される第1のキューとを含み、

前記第2の受信機が、

前記コマンド・ソースに結合される第2のバッファと、

前記第2のバッファからコマンドを受信するように結合

2

される第2の検証モジュールと、

前記第2の検証モジュールから検証済みコマンドを受信
 するように結合される第2のキューと、

を含む、請求項1記載のシステム。

【請求項6】前記第1のコマンド処理ユニットがアクテ
 イブ・ユニットとして指定され、前記第2のコマンド処
 理ユニットが待機ユニットとして指定されるとき、前記

第1及び前記第2のユニット間インタフェースが、
 前記第1のキューと前記第2のキューとの間の第1の動

10 作リンクと、

前記第1のキューと前記第2のバッファとの間の第2の
 動作リンクと、

前記第1の応答発生機と前記第2のキューとの間の第3
 の動作リンクと、

を含む、請求項5記載のシステム。

【請求項7】前記指定モジュールが、

前記第1のコマンド処理ユニットが故障を発生したことを
 検出する手段と、

前記第2のコマンド処理ユニットをアクティブ・ユニッ
 トとして再指定する手段と、

前記第1のコマンド処理ユニットが復元されたことを検
 出する手段と、

前記第1のコマンド処理ユニットを待機ユニットとして
 再指定する手段と、

を含む、請求項5記載のシステム。

【請求項8】前記第2のコマンド処理ユニットがアクテ
 イブ・ユニットとして指定され、前記第1のコマンド処
 理ユニットが待機ユニットとして指定されるとき、前記

第1及び前記第2のユニット間インタフェースが、
 前記第2のキューと前記第1のバッファとの間の第4の

30 動作リンクと、

前記第2の応答発生機と前記第1のキューとの間の第5
 の動作リンクと、

を含む、請求項7記載のシステム。

【請求項9】耐障害コマンド処理方法であって、

a) 第1のコマンド処理ユニットをアクティブ・ユニッ
 トとして指定し、第2のコマンド処理ユニットを待機ユ
 ニットとして指定するステップと、

b) コマンド・ソースにより送信されるコマンドを前記
 第1のコマンド処理ユニットのコマンド・バッファに受

信するステップと、
 c) コマンドを前記第2のコマンド処理ユニットのコマ
 ンド・バッファに受信しようと試みるステップと、

d) 前記第1のコマンド処理ユニットのコマンド検証モ
 ジュール内でコマンドを検証するステップと、

e) 前記第1のコマンド処理ユニットのコマンド・キュー
 にコマンドを待機するステップと、

f) キュー・メッセージを前記第1のコマンド処理ユニ
 ットから前記第2のコマンド処理ユニットへ送信するス

テップと、
 g) キュー・メッセージを前記第2のコマンド処理ユニ
 ットから前記第1のコマンド処理ユニットへ送信するス

テップと、
 h) キュー・メッセージを前記第1のコマンド処理ユニ
 ットから前記第2のコマンド処理ユニットへ送信するス

テップと、
 i) キュー・メッセージを前記第2のコマンド処理ユニ
 ットから前記第1のコマンド処理ユニットへ送信するス

テップと、
 j) キュー・メッセージを前記第1のコマンド処理ユニ
 ットから前記第2のコマンド処理ユニットへ送信するス

テップと、
 k) キュー・メッセージを前記第2のコマンド処理ユニ
 ットから前記第1のコマンド処理ユニットへ送信するス

50

g) 前記第2のコマンド処理ユニットの前記バッファ内においてコマンドが見出されない場合、前記第2のコマンド処理ユニットのコマンド・キューにコマンドを待機するステップと、

h) 前記第1のコマンド処理ユニットにコマンドにตอบสนองするように指令するステップと、

i) 前記第1のコマンド処理ユニットから前記第2のコマンド処理ユニットにデキュー・メッセージを送信するステップと、

j) 前記第2のコマンド処理ユニットの前記コマンド・キューからコマンドを待機解除するステップと、を含む方法。

【請求項10】前記コマンド待機解除ステップが、前記第2のコマンド処理ユニットがコマンドにตอบสนองすることなしに、該コマンドを待機解除するステップを含む、請求項9記載の方法。

【請求項11】識別子を前記コマンド・ソースから受信されたコマンドに関連付けるステップを含む、請求項9記載の方法。

【請求項12】前記キュー・メッセージ送信ステップが、前記コマンド識別子及びキュー指示を含むメッセージを送信するステップを含む、請求項11記載の方法。

【請求項13】前記キュー待機解除メッセージ送信ステップが、前記コマンド識別子及びキュー待機解除指示を含むメッセージを送信するステップを含む、請求項11記載の方法。

【請求項14】前記アクティブ・コマンド処理ユニットと前記待機コマンド処理ユニットとの間で、周期的にハンドシェーキング信号を交換するステップを含み、一方のユニットによる前記ハンドシェーキング信号の受信の失敗が、他のユニットにおいて故障指示となる、請求項9記載の方法。

【請求項15】前記第1のコマンド処理ユニットが故障した場合、

k) 前記第1のコマンド処理ユニットをオフラインにするステップと、

l) 前記第2のコマンド処理ユニットをアクティブ・ユニットとして再指定するステップと、

m) 前記第2のコマンド処理ユニットに、該第2のコマンド処理ユニットの前記コマンド・バッファ及び前記コマンド・キュー内のコマンドを処理するように指令するステップと、を含む、請求項9記載の方法。

【請求項16】n) 故障の修復後、前記第1のコマンド処理ユニットを待機ユニットとして再指定するステップを含む、請求項15記載の方法。

【請求項17】o) 前記ステップf)の後、前記第2のコマンド処理ユニットの前記コマンド・バッファ内にコマンドが存在するかどうかを判断するステップと、

p) 前記第2のコマンド処理ユニットの前記コマンド・

バッファ内にコマンドが存在しない場合、前記第2のコマンド処理ユニットから前記第1のコマンド処理ユニットへエラー・メッセージを送信するステップと、

q) 前記第1のコマンド処理ユニットの前記コマンド・キューから前記第2のコマンド処理ユニットの前記コマンド・キューにコマンドを送信するステップと、を含む、請求項15記載の方法。

【請求項18】前記キュー・メッセージ送信ステップが、前記コマンド識別子及びキュー指示を含むメッセージを送信するステップを含み、

前記第1のコマンド処理ユニットからコマンドを送信する前記ステップが、該コマンド及び関連識別子を送信するステップを含む、請求項17記載の方法。

【請求項19】耐障害自動化データ記憶及び検索ライブラリであって、

複数の取外し可能記憶媒体の選択された1つとの間でデータを読み書きするデータ・ドライブと、

前記データ・ドライブとホスト・プロセッサとの間に相互接続されるドライブ制御装置と、

格納セルと前記データ・ドライブとの間で選択媒体を移送する媒体アクセス機構と、

前記ドライブ制御装置及び前記媒体アクセス機構に結合される第1及び第2のライブラリ・マネージャとを含み、

前記第1及び前記第2のライブラリ・マネージャがそれぞれ、

前記ドライブ制御装置からコマンドを受信するように結合されるコマンド・バッファと、

前記コマンド・バッファからコマンドを受信するように結合されるコマンド検証モジュールと、

前記コマンド検証モジュールから検証済みコマンドを受信するように結合されるコマンド・キューと、

前記コマンド・キューから検証済みコマンドを受信し、前記ドライブ制御装置にコマンド応答を送信するように結合されるコマンド応答発生器と、

インタフェースであって、

前記第1のライブラリ・マネージャの前記コマンド・キューと前記第2のライブラリ・マネージャの前記コマンド・キューとの間の第1の動作リンクと、

前記第1のライブラリ・マネージャの前記コマンド・キューと、前記第2のライブラリ・マネージャの前記コマンド・バッファとの間の第2の動作リンクと、

前記第1のライブラリ・マネージャの前記コマンド応答発生器と、前記第2のライブラリ・マネージャの前記コマンド・キューとの間の第3の動作リンクと、

を含む、前記インタフェースと、を含む、ライブラリ。

【請求項20】最初に前記第1のライブラリ・マネージャをアクティブ・ユニットとして、前記第2のライブラリ・マネージャを待機ユニットとして指定する手段を含

む、請求項19記載のライブラリ。

【請求項21】前記第1のライブラリ・マネージャが、コマンドが検証され、前記第1のライブラリ・マネージャの前記コマンド・キューに転送された後、前記第2のライブラリ・マネージャの前記コマンド・バッファから、前記第2のライブラリ・マネージャの前記コマンド・キューに前記コマンドを転送するように指令する手段と、

前記第1のライブラリ・マネージャの前記コマンド応答発生器が前記コマンドに対する応答を完了後、前記第2のライブラリ・マネージャの前記コマンド・キューから前記コマンドを待機解除するように指令する手段と、を含む、請求項19記載のライブラリ。

【請求項22】前記第1のライブラリ・マネージャが、コマンドが前記第2のライブラリ・マネージャの前記コマンド・バッファに受信されていない場合、前記コマンドが検証され、前記第1のライブラリ・マネージャの前記コマンド・キューに転送された後、前記第1のライブラリ・マネージャの前記コマンド・キューから、前記第2のライブラリ・マネージャの前記コマンド・キューに前記コマンドを転送するように指令する手段と、

前記第1のライブラリ・マネージャの前記コマンド応答発生器が前記コマンドに対する応答を完了後、前記第2のライブラリ・マネージャの前記コマンド・キューから前記コマンドを待機解除するように指令する手段とを含む、請求項20記載のライブラリ。

【請求項23】前記第1のライブラリ・マネージャが故障を発生した場合、前記第2のライブラリ・マネージャを待機ユニットとして再指定し、前記故障が修復したとき、前記第1のライブラリ・マネージャを待機ユニットとして再指定する手段を含む、請求項20記載のライブラリ。

【請求項24】前記インタフェースが、前記第2のライブラリ・マネージャの前記コマンド・キューと、前記第1のライブラリ・マネージャの前記コマンド・バッファとの間の第4の動作リンクと、前記第2のライブラリ・マネージャの前記コマンド応答発生器と、前記第1のライブラリ・マネージャの前記コマンド・キューとの間の第5の動作リンクと、を含む、請求項23記載のライブラリ。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は冗長コンポーネントを有する耐障害（フォールト・トレラント）システムに関し、特に、冗長プロセッサにより、従来の耐障害処理において要求されるオーバーヘッドを低減する方法及び装置に関する。

【0002】

【従来の技術】“致命的”エラーが発生した場合にもデータ処理の継続を保证するために、しばらくの間、大規模

コンピュータの導入において耐障害コンピュータ・システムが使用されてきた。こうしたシステムは、一方のユニットが故障した場合に、他方が即時に引継ぐように、同時且つ並列に動作する冗長プロセッサ及び他のコンポーネントを含む。一方のユニットに伝送される全てのコマンド及びデータは、他方へも伝送されなければならないことが明らかであろう。更に、一方のユニットにより生成される全ての応答は、他のユニットにより生成される応答と一致しなければならない。応答が異なると、エラーまたは故障が指摘され、故障箇所及び実行されるべき適切な修正作業を決定するために、診断プロシージャが実行されなければならない。こうした修正作業には、故障ユニットをオフラインにし、他のユニットをオンラインにすることが含まれる。冗長システムにおいて、全てのコマンドが同時に受信され、処理されることを保証するために、通常、拡張ユニット間通信が必要とされる。こうした通信は、多大なオーバーヘッド負荷をユニット及びトータル・システムの性能に課することになり、システムのコストを増加させる。結果的に、耐障害冗長システムは、ハイ・エンドのクリティカル・アプリケーションにおいて一般的に使用されてきた。

【0003】しかしながら、パワー・コンピュータ技術の価格の低下により、耐障害システムはより現実的となった。例えばIBMにより開発された3495テープ・ライブラリ・データサーバなどの自動化データ記憶及び検索ライブラリ・システムにおいて、ライブラリ内の全ての個々のコンポーネントの機能を調整するライブラリ・マネージャは、従来のパーソナル・コンピュータである。そのコスト自身は、ライブラリの総コストの比較的小さな部分を占めるため、耐障害冗長ライブラリ・マネージャは、多くの顧客により所望される経済的に実現可能なオプションになった。しかしながら、今日まで拡張ユニット間通信に対する要求は依然として減っていない。

【0004】

【発明が解決しようとする課題】上述を鑑み、本発明の目的は、従来システムよりも効率的なユニット間通信を有する耐障害冗長処理システムを提供することである。

【0005】本発明の別の目的は、バックアップ・ユニットの処理負荷が低減されるこうしたシステムを提供することである。

【0006】更に本発明の別の目的は、自動化データ記憶及び検索システムにおける耐障害ライブラリ・マネージャを提供することである。

【0007】

【課題を解決するための手段】これらの目的が、本発明において冗長コマンド・プロセッサを使用する耐障害システムを提供することにより達成される。冗長コマンド・プロセッサは、一方のプロセッサの故障にも関わらず、受信されたものの有効と見なされないコマンドを有

効化し、実行のためにキューに待機されたコマンドを失わないようにし、また実行中のプロセスのコマンドを完了して、まだ送信されていない応答を送信することを保証する。更に冗長ユニット間の同期が、低減されたユニット間通信により維持される。

【0008】より詳細には、アクティブ・ユニットと待機ユニットの両方によりコマンドが受信され、キューに待機されるが、コマンドはアクティブ・ユニットによってのみ処理される。アクティブ・ユニットが故障した場合には、待機ユニットが自動的にオンラインにされ、未

処理のコマンドを処理し、応答をコマンド・ソースに送信する。故障ユニットが復元すると、そのユニットが待機ユニットとなる。

【0009】一実施例では、耐障害コマンド処理のためのシステムが提供され、このシステムはコマンド・ソースからコマンドを受信する第1及び第2の受信機と、第1及び第2の受信機にそれぞれ相互接続され、応答をコマンド・ソースに送信する第1及び第2の送信機と、第1の受信機及び第1の送信機に相互接続される第1のユニット間インタフェースと、第2の受信機、第2の送信機及び第1のユニット間インタフェースに相互接続される第2のユニット間インタフェースとを含む。第1の受信機、第1の送信機及び第1のユニット間インタフェースは、第1のコマンド処理ユニットを構成し、第2の受信機、第2の送信機及び第2のユニット間インタフェースは、第2のコマンド処理ユニットを構成する。システムは更に、第1のコマンド処理ユニットをアクティブ・ユニットとして、また第2のコマンド処理ユニットを待機ユニットとして指定する指定モジュールを含み、それにより第1の受信機及び第1の送信機が第2の受信機に動作的にリンクされる。指定モジュールは更に、第1のコマンド処理ユニットが復元されたことを検出する手段と、第1のコマンド処理ユニットを待機ユニットとして再指定する手段とを含み、それにより第2の受信機及び第2の送信機が第1の受信機に動作的にリンクされる。

【0010】

【実施例】図1は、コマンド・ソース16に相互接続される冗長コマンド・プロセッサA及びB（それぞれ12及び14）を使用する本発明の耐障害システム10のブロック図である。コマンド・プロセッサ12及び14は冗長ライブラリ・マネージャであり、コマンド・ソース16は自動化データ記憶及び検索ライブラリ・システム（図3に関連して後述される）内のドライブ制御装置である。しかしながら、本発明はこの特定の構成に限るものではなく、他の動作環境にも同様に適応可能である。各コマンド・プロセッサ・ユニット12及び14は、ユニット制御装置121及び141、コマンド・バッファ122及び142、コマンド検証モジュール124及び144、コマンド・キュー126及び146、及びコマンド応答発生器128及び148を含む。各ユニット1

2及び14はまた、簡略化のために図1では省略されて示されない他のモジュール（図4に関連して後述される）も含む。

【0011】各ユニット12及び14は、コマンド・ソース16からコマンドを受信するために、またコマンド応答を送信するために、相互接続される。更にユニット間通信が、2つのユニット12及び14のコマンド・キュー126と146との間、及び一方のユニット12または14のコマンド・キュー126または146と、他方のユニット14または12のコマンド・バッファ142または122との間、更に一方のユニット12または14のコマンド・キュー126または146と、他方のユニット14または12のコマンド応答バッファ148または128との間のリンクにより提供される。これらのリンクは、実際にはユニット間通信を指令するユニット制御装置121及び141を介して処理される。

【0012】コマンド・ソース16から伝送される各コマンドは、関連する識別子を有する。一実施例では識別子は4バイト長である。特定のコマンドのステータスを問い合わせまたは提供する短いメッセージが、2つのユニット12及び14の間で伝送される。更に周期的なハンドシェーキングが、2つのユニット12及び14の間で発生する。すなわち、一方のユニット制御装置121または141が他のユニット制御装置141または121に短い信号を送信し、後者が肯定応答を返却する。一方のユニット制御装置121または141が適切な信号の受信を失敗すると、これは他のユニット14または12が問題を被ったことを示す。結果的に、ユニット間通信の数及びサイズの両方が多大に低減される。

【0013】耐障害システム10の動作について、図2を参照しながら説明する。起動（ステップ200）の後、2つのコマンド・プロセッサ・ユニット12または14の一方がアクティブ・ユニットとして、他方が待機ユニットとして指定される（ステップ202）。ここではユニットA12が最初にアクティブ・ユニットとして指定され、ユニットB14が待機ユニットとして指定されるものとする。この構成では、アクティブ・ユニット12のコマンド・キュー126は、機能的に待機ユニット14のコマンド・バッファ142及びコマンド・キュー146にリンクされ、アクティブ・ユニット12のコマンド応答発生器128は、待機ユニット14のコマンド・キュー146にリンクされる。上述のように、こうしたリンクはユニット制御装置121及び141により管理される。両ユニット12及び14のコマンド・バッファ122及び142がコマンド・ソース16に結合され、アクティブ・ユニット12のコマンド応答発生器128もコマンド・ソース16に結合される。

【0014】両コマンド・バッファ122及び142は、コマンド・ソース16によりコマンドが送信されるのを待機する（ステップ204）。コマンドはアクティ

ブ・ユニットA12により受信され(ステップ206)、コマンド・バッファ122に配置される。コマンドが待機ユニットB14によっても受信されると(ステップ208)、そのコマンドは待機コマンド・バッファ142に配置される(ステップ210)。コマンドがアクティブ・コマンド・バッファ122に配置されると(ステップ206)、コマンドはアクティブ・コマンド検証器124により妥当性検証され(ステップ212)、次にアクティブ・コマンド・キュー126に転送される。コマンドを待機バッファ142から待機コマンド・キュー146に転送するように指示するメッセージが、アクティブ・ユニット制御装置121により、待機ユニット制御装置141に送信される(ステップ216)。妥当性検証は待機ユニット14によつては実行されない。ステップ212でのアクティブ・ユニットA12による妥当性検証が不成功の場合、エラー回復プロシージャが呼出される(一般にはコマンド・ソース16によるコマンドの再送の要求を含む)。妥当性検証が成功の場合には、待機ユニット14は妥当性検証を繰返す必要はない。

【0015】次に、コマンドが待機バッファ142において成功裡に見出されたかどうかの判断チェックが実行される(ステップ218)。否定の場合、エラー・メッセージがアクティブ・ユニット制御装置121に送信され(ステップ220)、アクティブ・ユニット制御装置121はそれに応答して、コマンド・キュー126から待機バッファ142にコマンドを送信する(ステップ222)。次に、或いはステップ218でコマンドが待機バッファ142内で見出されない場合、コマンドが待機バッファ142から待機キュー146に転送される(ステップ224)。

【0016】ハードウェア故障または通信故障は、コマンド処理の間の任意の時刻において発生する可能性があり、アクティブ及び待機ユニット制御装置121及び141は、周期的に短いハンドシェーキング信号をユニット間リンクを横断して交換する。ユニット12または14の一方が故障を経験するか、信号の交換が途切れると、他のユニットの他のユニット制御装置が故障を検出し、エラー回復プロシージャを開始する。ステップ226は、故障判断が実行されるプロシージャにおける多くのステージの一例であり、この特定のステージにおける一例として示されているだけで、これに限るものではない。アクティブ・ユニットA12が故障に遭遇しない限り(たとえ待機ユニットB14が故障したとしても)、アクティブ・ユニットA12はコマンドの処理を継続する。このように、コマンドが待機ユニットB14のキューに待機された後、アクティブ・コマンド応答発生器128がコマンドに対する適切な応答を発生し(ステップ228)、それをコマンド・ソース16に送信する。コマンド・ソース16からのコマンドは、アクティブ・ユ

ニットA12により開始される特定のアクションに対する要求を含むことが理解されよう。アクティブ・ユニットA12はこうしたアクションをステップ228の一部として開始し、コマンド・ソース16に送信される適切な応答は、アクションが開始されたことを示す。応答を受信すると、アクティブ・ユニット制御装置121は待機ユニット制御装置141に、待機キュー146からコマンドを除去するように指令する。すなわち、待機ユニットB14は自身のコマンド応答を生成する必要はない。プロシージャは次にステップ204に戻り、別のコマンドを待機する。すなわち、コマンドがアクティブ・コマンド・バッファ122に受信されると、別のコマンドがコマンド・ソースにより送信されることが理解されよう。このようにして、上述の処理が継続され、多くのコマンドが一度に"パイプライン"形式で転送される。

【0017】アクティブ・ユニットA12が故障すると(ステップ226)、待機ユニットB14が活動化され(ステップ234)、アクティブ・ユニットとして再指定され(ステップ236)、故障した以前のアクティブ・ユニットA12は適切なサービスまたは回復のためにオフライン化される。以前の待機バッファB142内のコマンドが、アクティブ・ユニットA12に関連して上述されたように妥当性検証され、キューに待機され(ステップ238)、応答が生成され(ステップ240)、コマンド・ソース16に送信される。バッファまたはキュー内の他のコマンドについても処理される。次にフローの最初に復帰し、更にコマンドを待機する。故障したユニットA12が修復され、オンラインにされると、ユニット間ハンドシェーキングが再開し、ユニットA12は待機ユニットとして指定される。

【0018】図3は、本発明が組込まれる自動化データ記憶及び検索ライブラリ・システム30のブロック図である。ライブラリ30は相互接続される1つまたは複数のホスト・コンピュータ32に対する大容量記憶装置として機能する。ライブラリ30は、磁気テープ・カートリッジまたは光ディスク・カートリッジなどのデータ記憶媒体を格納するセルを含む。セルの数はライブラリ30の構成に依存するが、数個乃至数万個の範囲に及ぶ。ライブラリ30はまた、媒体との間でデータを読み書きするための1つまたは複数のデータ・ドライブ34を含む。1つまたは複数のドライブ34はドライブ制御装置36と相互接続され、各ドライブ制御装置36はホスト32と相互接続される。

【0019】ドライブ制御装置36はまた、ライブラリ30の物理作業をモニタ及び管理するライブラリ・マネージャ40と相互接続される。ライブラリ・マネージャ40はオペレータ端末38及びアクセス制御装置50に相互接続される。アクセス制御装置50は、媒体を格納セルとドライブ34との間で移動するロボット移送機構51、媒体を掴むためにロボット51上に搭載されるグ

リップ52、媒体上の識別ラベルを読取るビジョン・システム54、媒体をオペレータとの間で受渡しする入出力ステーション56、及びライブラリ・アクセス・パネルがオープンされた時を示すインターロック・センサ58をモニタ及び管理する。

【0020】オペレーションにおいて、ホスト32はデータ・アクセス要求をドライブ制御装置36に送信する。制御装置36は要求をライブラリ・マネージャ40に送信する。ライブラリ・マネージャ40内のデータベースは、アクセスされる媒体の識別及びロケーションを示し、ライブラリ・マネージャ40は媒体をその現在位置からドライブ34の1つに移送するようにコマンドを発行する。要求された媒体がドライブ34の1つにロードされると、ホスト32は適切なドライブ制御装置36を介して通知され、データが媒体との間で転送される。

【0021】図4はライブラリ・マネージャ40の機能ブロック図であり、ライブラリ・マネージャ40によりコマンドが処理される様子を説明するために参照される。コマンド処理が特にライブラリ・マネージャ40に関連して述べられるが、他のアプリケーションにおいても使用される。第1のインタフェース402は、ライブラリ・マネージャ40をドライブ制御装置36と相互接続する。第2のインタフェース404は、ライブラリ・マネージャ40をオペレータ端末38及びアクセス制御装置50と相互接続する。コマンドが第1のインタフェース402を介してホスト・メッセージ・ハンドラ406に受信され、次にそのコマンドが有効なコマンドであり、適切な形式であることを保証するために、コマンド検証器408にパスされる。コマンドが妥当であったかどうかを示す応答が、応答発生器410により生成され、出力ホスト・メッセージ・ハンドラ412に送信され、第1のインタフェース402を介してドライブ制御装置36に出力される。

【0022】コマンドが成功裡に検証されると、そのコマンドはコマンド・キュー・マネージャ414内の以前に受信されて検証されたコマンドと一緒にキューに配置される。コマンドが続いて実行のためにキューからフェッチされる。すなわち、アクセス制御装置50またはオペレータ端末38への適切なコマンドが、コマンド発生器416により生成され、実行コーディネータ418及び第2のインタフェース404を介して送信出力される。コマンドが実行されるとオペレータ端末38またはアクセス制御装置50が、完了信号を第2のインタフェース404を介して応答ハンドラ420に送信し、完了信号は更にコマンド発生器416に伝達される。その後、コマンド発生器416は応答発生器410に、コマンドが実行されたことを信号で知らせる。するとコマンド発生器410は、出力ホスト・メッセージ・ハンドラ412及び第1のインタフェース402を介して、ドライブ制御装置36に、コマンドが完了したことを示すメ

ッセージを送信する。

【0023】ライブラリ・マネージャ40はまた、ルーチン・エラー、手動オペレーション、及び媒体データベースを管理する適切なモジュール（それぞれ要素422、424、426）を有する。更にコマンド検証器408、コマンド・キュー・マネージャ414、及び応答発生器410が、動作的にライブラリ・マネージャ40の様々な他の局面に相互接続される。こうしたものには、例えば自動化カートリッジ・ロード428、入出力マネージャ430、在庫マネージャ432及び自動クリーン機能434（適切な時刻においてドライブ・クリーニングを開始及びモニタする）が含まれる。動作状態についても別のモジュール436によりモニタされ、適切なステータス・メッセージがドライブ制御装置36に送信される。

【0024】ライブラリ・マネージャ40の任意のサブシステムが故障すると、修復が達成されるまで、全体のライブラリ30が遮断されなければならないかもしれない。任意の期間ダウン状態のライブラリまたは任意の他のコンピュータ・システムを有することは、費用を費やす上にユーザにとって腹立たしいことである。結局、本発明は冗長コマンド・マネージャを使用する耐障害システムを提供する。しかしながら、両方のコマンド・プロセッサによりコマンドが同時に処理される従来の冗長システムとは異なり、本発明のシステムはユニット間オーバヘッドを低減し、待機ユニットが活動化されるまで、待機ユニットによるコマンド処理を排除する。

【0025】本発明は特定の実施例について述べられてきたが、当業者には本発明の精神及び範囲を逸脱することなしに、その形態及び詳細における様々な変更が可能であることが理解されよう。

【0026】まとめとして、本発明の構成に関して以下の事項を開示する。

【0027】(1) 耐障害コマンド処理のためのシステムであって、コマンド・ソースからコマンドを受信する第1の受信機と、前記第1の受信機に結合され、前記コマンド・ソースに応答を送信する第1の送信機と、前記第1の受信機及び前記第1の送信機に結合される第1のユニット間インタフェースと、を含む、第1のコマンド処理ユニットと、前記コマンド・ソースからコマンドを受信する第2の受信機と、前記第2の受信機に結合され、前記コマンド・ソースに応答を送信する第2の送信機と、前記第2の受信機、前記第2の送信機及び前記第1のユニット間インタフェースに結合される第2のユニット間インタフェースと、を含む、第2のコマンド処理ユニットと、前記第1のコマンド処理ユニットをアクティブ・ユニットとして指定し、前記第2のコマンド処理ユニットを待機ユニットとして指定する指定モジュールと、を含み、前記第1の受信機及び前記第1の送信機が前記第2の受信機に動作的にリンクされる、システム。

(2) コマンド・ソースからのコマンドが第2の受信機により受信されない場合に、該コマンドを前記第1の受信機から前記第2の受信機に送信するように指令する手段を含む、前記(1)記載のシステム。

(3) 前記指定モジュールが、前記第1のコマンド処理ユニットが故障を発生したことを検出する手段と、前記第2のコマンド処理ユニットをアクティブ・ユニットとして再指定する手段と、を含む、前記(1)記載のシステム。

(4) 前記指定モジュールが、前記第1のコマンド処理ユニットが復元されたことを検出する手段と、前記第1のコマンド処理ユニットを待機ユニットとして再指定する手段とを含む、前記第2の受信機及び前記第2の送信機が動作的に前記第1の受信機にリンクされる、前記(3)記載のシステム。

(5) 前記第1の受信機が、前記コマンド・ソースに結合される第1のバッファと、前記第1のバッファからコマンドを受信するように結合される第1の検証モジュールと、前記第1の検証モジュールから検証済みコマンドを受信するように結合される第1のキューとを含む、前記第2の受信機が、前記コマンド・ソースに結合される第2のバッファと、前記第2のバッファからコマンドを受信するように結合される第2の検証モジュールと、前記第2の検証モジュールから検証済みコマンドを受信するように結合される第2のキューと、を含む、前記(1)記載のシステム。

(6) 前記第1のコマンド処理ユニットがアクティブ・ユニットとして指定され、前記第2のコマンド処理ユニットが待機ユニットとして指定されるとき、前記第1及び前記第2のユニット間インタフェースが、前記第1のキューと前記第2のキューとの間の第1の動作リンクと、前記第1のキューと前記第2のバッファとの間の第2の動作リンクと、前記第1の応答発生機と前記第2のキューとの間の第3の動作リンクと、を含む、前記(5)記載のシステム。

(7) 前記指定モジュールが、前記第1のコマンド処理ユニットが故障を発生したことを検出する手段と、前記第2のコマンド処理ユニットをアクティブ・ユニットとして再指定する手段と、前記第1のコマンド処理ユニットが復元されたことを検出する手段と、前記第1のコマンド処理ユニットを待機ユニットとして再指定する手段と、を含む、前記(5)記載のシステム。

(8) 前記第2のコマンド処理ユニットがアクティブ・ユニットとして指定され、前記第1のコマンド処理ユニットが待機ユニットとして指定されるとき、前記第1及び前記第2のユニット間インタフェースが、前記第2のキューと前記第1のバッファとの間の第4の動作リンクと、前記第2の応答発生機と前記第1のキューとの間の第5の動作リンクと、を含む、前記(7)記載のシステム。

(9) 耐障害コマンド処理方法であって、

- a) 第1のコマンド処理ユニットをアクティブ・ユニットとして指定し、第2のコマンド処理ユニットを待機ユニットとして指定するステップと、
- b) コマンド・ソースにより送信されるコマンドを前記第1のコマンド処理ユニットのコマンド・バッファに受信するステップと、
- c) コマンドを前記第2のコマンド処理ユニットのコマンド・バッファに受信しようと試みるステップと、
- d) 前記第1のコマンド処理ユニットのコマンド検証モジュール内でコマンドを検証するステップと、
- e) 前記第1のコマンド処理ユニットのコマンド・キューにコマンドを待機するステップと、
- f) キュー・メッセージを前記第1のコマンド処理ユニットから前記第2のコマンド処理ユニットへ送信するステップと、
- g) 前記第2のコマンド処理ユニットの前記バッファ内においてコマンドが見出されない場合、前記第2のコマンド処理ユニットのコマンド・キューにコマンドを待機するステップと、
- h) 前記第1のコマンド処理ユニットにコマンドに応答するように指令するステップと、
- i) 前記第1のコマンド処理ユニットから前記第2のコマンド処理ユニットにデキュー・メッセージを送信するステップと、
- j) 前記第2のコマンド処理ユニットの前記コマンド・キューからコマンドを待機解除するステップと、を含む方法。

(10) 前記コマンド待機解除ステップが、前記第2のコマンド処理ユニットがコマンドに応答することなしに、該コマンドを待機解除するステップを含む、前記(9)記載の方法。

(11) 識別子を前記コマンド・ソースから受信されたコマンドに関連付けるステップを含む、前記(9)記載の方法。

(12) 前記キュー・メッセージ送信ステップが、前記コマンド識別子及びキュー指示を含むメッセージを送信するステップを含む、前記(11)記載の方法。

(13) 前記キュー待機解除メッセージ送信ステップが、前記コマンド識別子及びキュー待機解除指示を含むメッセージを送信するステップを含む、前記(11)記載の方法。

(14) 前記アクティブ・コマンド処理ユニットと前記待機コマンド処理ユニットとの間で、周期的にハンドシェーキング信号を交換するステップを含み、一方のユニットによる前記ハンドシェーキング信号の受信の失敗が、他のユニットにおいて故障指示となる、前記(9)記載の方法。

(15) 前記第1のコマンド処理ユニットが故障した場合、k) 前記第1のコマンド処理ユニットをオフライン

15

にするステップと、l) 前記第2のコマンド処理ユニットをアクティブ・ユニットとして再指定するステップと、m) 前記第2のコマンド処理ユニットに、該第2のコマンド処理ユニットの前記コマンド・バッファ及び前記コマンド・キュー内のコマンドを処理するように指令するステップと、を含む、前記(9)記載の方法。

(16) n) 故障の修復後、前記第1のコマンド処理ユニットを待機ユニットとして再指定するステップを含む、前記(15)記載の方法。

(17) o) 前記ステップf)の後、前記第2のコマンド処理ユニットの前記コマンド・バッファ内にコマンドが存在するかどうかを判断するステップと、

p) 前記第2のコマンド処理ユニットの前記コマンド・バッファ内にコマンドが存在しない場合、前記第2のコマンド処理ユニットから前記第1のコマンド処理ユニットへエラー・メッセージを送信するステップと、

q) 前記第1のコマンド処理ユニットの前記コマンド・キューから前記第2のコマンド処理ユニットの前記コマンド・キューにコマンドを送信するステップと、を含む、前記(15)記載の方法。

(18) 前記キュー・メッセージ送信ステップが、前記コマンド識別子及びキュー指示を含むメッセージを送信するステップを含み、前記第1のコマンド処理ユニットからコマンドを送信する前記ステップが、該コマンド及び関連識別子を送信するステップを含む、前記(17)記載の方法。

(19) 耐障害自動化データ記憶及び検索ライブラリであって、複数の取外し可能記憶媒体の選択された1つとの間でデータを読み書きするデータ・ドライブと、前記データ・ドライブとホスト・プロセッサとの間に相互接続されるドライブ制御装置と、格納セルと前記データ・ドライブとの間で選択媒体を移送する媒体アクセス機構と、前記ドライブ制御装置及び前記媒体アクセス機構に結合される第1及び第2のライブラリ・マネージャとを含み、前記第1及び前記第2のライブラリ・マネージャがそれぞれ、前記ドライブ制御装置からコマンドを受信するように結合されるコマンド・バッファと、前記コマンド・バッファからコマンドを受信するように結合されるコマンド検証モジュールと、前記コマンド検証モジュールから検証済みコマンドを受信するように結合されるコマンド・キューと、前記コマンド・キューから検証済みコマンドを受信し、前記ドライブ制御装置にコマンド応答を送信するように結合されるコマンド応答発生器と、インタフェースであって、前記第1のライブラリ・マネージャの前記コマンド・キューと前記第2のライブラリ・マネージャの前記コマンド・バッファとの間の第2の動作リンクと、前記第1のライブラリ・マネージャの前記コマンド・キューと、前記第2のライブラリ・マネージャの前記コマンド・バッファとの間の第3の動作リンクと、を含む、前記(19)記載のライブラリ。

16

ンド応答発生器と、前記第2のライブラリ・マネージャの前記コマンド・キューとの間の第3の動作リンクと、を含む、前記インタフェースと、を含む、ライブラリ。

(20) 最初に前記第1のライブラリ・マネージャをアクティブ・ユニットとして、前記第2のライブラリ・マネージャを待機ユニットとして指定する手段を含む、前記(19)記載のライブラリ。

(21) 前記第1のライブラリ・マネージャが、コマンドが検証され、前記第1のライブラリ・マネージャの前記コマンド・キューに転送された後、前記第2のライブラリ・マネージャの前記コマンド・バッファから、前記第2のライブラリ・マネージャの前記コマンド・キューに前記コマンドを転送するように指令する手段と、前記第1のライブラリ・マネージャの前記コマンド応答発生器が前記コマンドに対する応答を完了後、前記第2のライブラリ・マネージャの前記コマンド・キューから前記コマンドを待機解除するように指令する手段と、を含む、前記(19)記載のライブラリ。

(22) 前記第1のライブラリ・マネージャが、コマンドが前記第2のライブラリ・マネージャの前記コマンド・バッファに受信されていない場合、前記コマンドが検証され、前記第1のライブラリ・マネージャの前記コマンド・キューに転送された後、前記第1のライブラリ・マネージャの前記コマンド・キューから、前記第2のライブラリ・マネージャの前記コマンド・キューに前記コマンドを転送するように指令する手段と、前記第1のライブラリ・マネージャの前記コマンド応答発生器が前記コマンドに対する応答を完了後、前記第2のライブラリ・マネージャの前記コマンド・キューから前記コマンドを待機解除するように指令する手段とを含む、前記(20)記載のライブラリ。

(23) 前記第1のライブラリ・マネージャが故障が発生した場合、前記第2のライブラリ・マネージャを待機ユニットとして再指定し、前記故障が修復したとき、前記第1のライブラリ・マネージャを待機ユニットとして再指定する手段を含む、前記(20)記載のライブラリ。

(24) 前記インタフェースが、前記第2のライブラリ・マネージャの前記コマンド・キューと、前記第1のライブラリ・マネージャの前記コマンド・バッファとの間の第4の動作リンクと、前記第2のライブラリ・マネージャの前記コマンド応答発生器と、前記第1のライブラリ・マネージャの前記コマンド・キューとの間の第5の動作リンクと、を含む、前記(23)記載のライブラリ。

【0028】

【発明の効果】以上説明したように、本発明によれば、従来システムよりも効率的なユニット間通信を有する耐障害冗長処理システムを提供することができる。

【0029】また、バックアップ・ユニットの処理負荷

17

が低減されるこうしたシステムを提供することができる。

【0030】更に、自動データ記憶及び検索システムにおける耐障害ライブラリ・マネージャを提供することができる。

【図面の簡単な説明】

【図1】本発明の耐障害システムの冗長ユニットの特定の要素及びそれらの間のユニット間動作リンクを表すブロック図である。

【図2】本発明の方法の一実施例の流れ図である。

【図3】本発明の耐障害システムが組込まれる自動データ記憶及び検索ライブラリ・システムのブロック図である。

【図4】本発明において使用されるライブラリ・マネージャの機能ブロック図である。

【符号の説明】

10 耐障害システム
12 冗長コマンド・プロセッサA
14 冗長コマンド・プロセッサB
16 コマンド・ソース
30 検索ライブラリ・システム
32 ホスト・コンピュータ
34 データ・ドライブ
36 ドライブ制御装置
38 オペレータ端末
40 ライブラリ・マネージャ
50 アクセス制御装置
51 ロボット移送機構

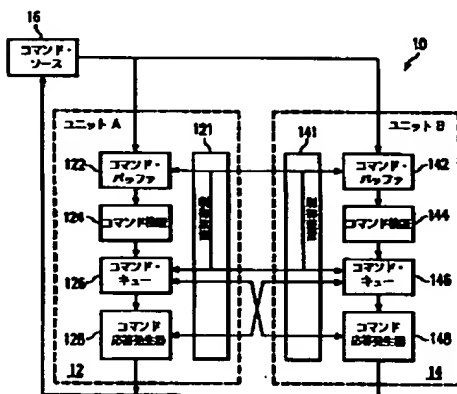
10

20

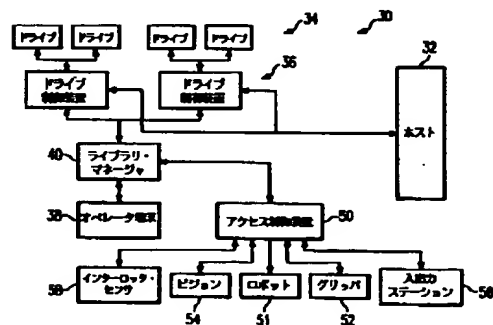
52 グリッパ
56 入出力ステーション
58 インターロック・センサ
121 ユニット制御装置
122 コマンド・バッファ
124 コマンド検証モジュール
126 コマンド・キュー
128 コマンド応答発生器
141 ユニット制御装置
142 コマンド・バッファ
144 コマンド検証モジュール
146 コマンド・キュー
148 コマンド応答発生器
402 第1のインタフェース
404 第2のインタフェース
406 ホスト・メッセージ・ハンドラ
408 コマンド検証器
410 応答発生器
412 出力ホスト・メッセージ・ハンドラ
414 コマンド・キュー・マネージャ
416 コマンド発生器
418 実行コーディネータ
420 応答ハンドラ
428 自動化カートリッジ・ローダ
430 入出力マネージャ
432 在庫マネージャ
434 自動クリーン機能
436 モジュール

18

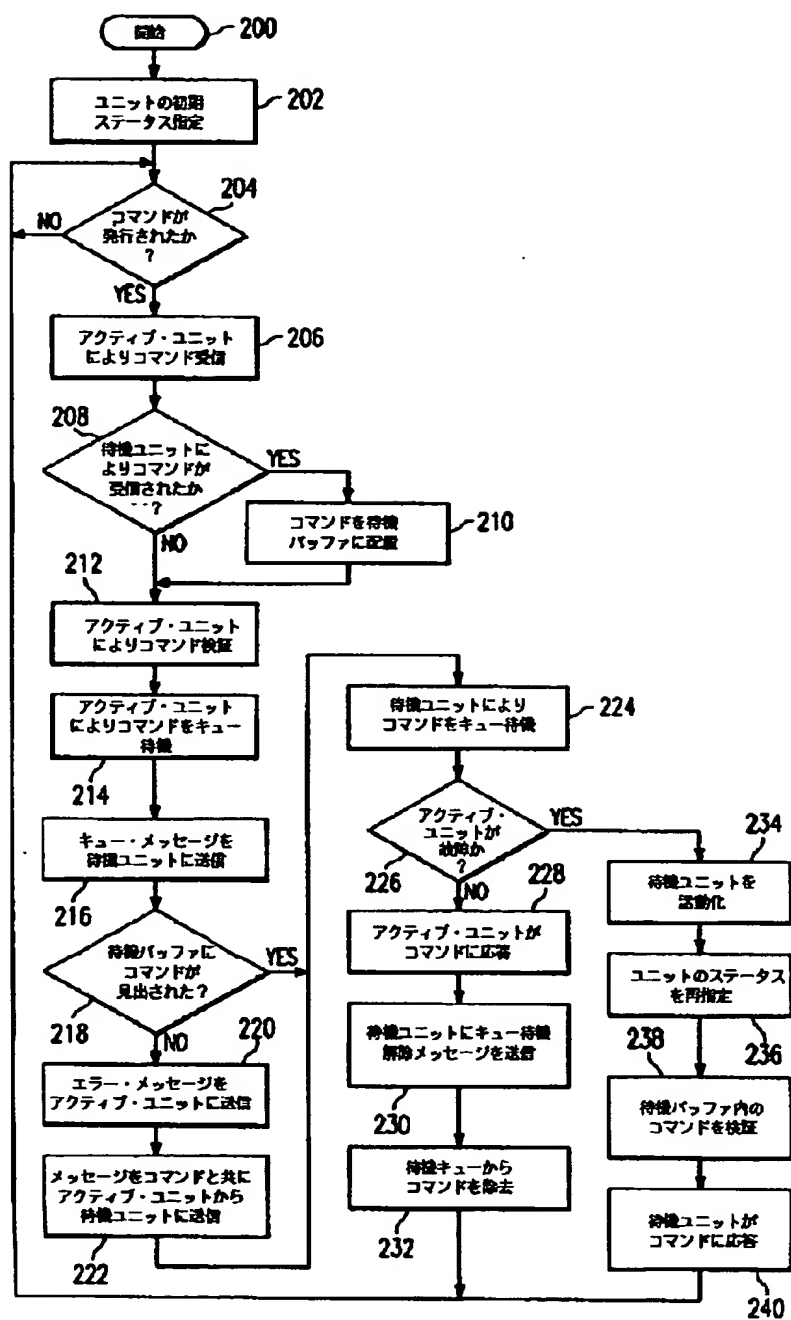
【図1】



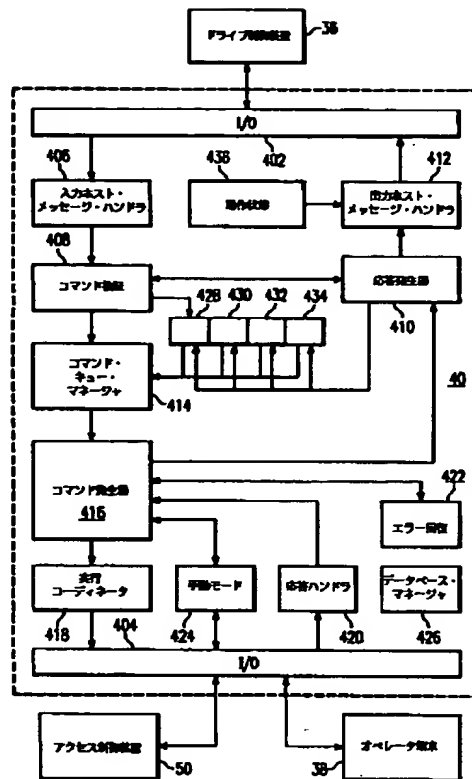
【図3】



【図2】



【図4】



フロントページの続き

(72)発明者 メリル・バーンズ・グレコ
 アメリカ合衆国85715、アリゾナ州ツーソ
 ン、イースト・ノルウッド・サークル
 7672

(72)発明者 マイケル・ジェームズ・ミドウズ
 アメリカ合衆国85715、アリゾナ州ツーソ
 ン、ノース・サビノ・キャニオン・ロード
 ナンバー1182 5051